

Simulated Annealing Optimization for Hydrocarbon Pipeline Networks

Diego A. Rodríguez,^{†,§,⊥} Paola P. Oteiza,^{*,‡,§,⊥} and Nélide B. Brignole^{§,⊥}

[†]Facultad de Ciencias Exactas, Universidad Nacional de Salta (UNSa), Av. Bolivia 5150, Salta, Argentina

[‡]Comisión de Investigaciones Científicas (CIC) Provincia de Buenos Aires, Calle 526 entre 10 y 11 CP1900 La Plata, Argentina

[§]Laboratorio de Investigación y Desarrollo en Computación Científica (LIDeCC), Departamento de Ciencias e Ingeniería de la Computación (DCIC), Universidad Nacional del Sur (UNS), Av. Alem 1253, Bahía Blanca, Argentina

[⊥]Planta Piloto de Ingeniería Química (PLAPIQUI) Complejo CCT-UAT, UNS-CONICET Camino La Carrindanga Km. 7, 8000 Bahía Blanca, Argentina

ABSTRACT: In this work the determination of optimally located pipeline networks has been proposed by means of the implementation of a metaheuristic algorithm called Simulated Annealing with GAMS (SAG) in order to find the best pipeline layout together with a subset of locations to install concentrating nodes. The strategy essentially consists of a hybridization of Simulated Annealing, combined with the well-known GAMS package. In particular, the sample cases consisted of finding the most convenient routes so as to transport natural gasoline from Santa Cruz (Argentina) gas fields to the processing plants. The SAG algorithm behaved satisfactorily because it proved to be efficient and flexible.

1. INTRODUCTION

The combinatorial problems are efficiently solved by exact algorithms, but as the problem size augments, the complexity increases exponentially. The metaheuristic techniques arise as alternative optimization methods to achieve good-quality solutions in relatively low computational times. Metaheuristic algorithms are iterative procedures for general-purpose searches that guide a subordinate heuristics by combining cleverly various concepts to explore and adequately exploit the search space. There is a wide variety of metaheuristics as a function of concepts such as the use of memory, the employment of a population of solutions, and the considered amount of neighborhoods.

Problem knowledge is always included in these search algorithms. In particular there are some heuristics based on trajectories, such as Simulated Annealing (SA), that carry out a local study about the search space. They analyze the present solution's neighborhood in order to decide how to continue with the search. In other words, they start from an initial solution and try to replace it iteratively by another solution of better quality located in the neighborhood. SA allows movements that worsen the present solution so as to run away from local optima.

SA is one of the most promising optimization techniques when an NP-hard problem¹ is to be solved. For these problems, practical users are basically content with a quasi-optimal solution with a reasonable computational cost.² The SA approach offers an unusual perspective toward optimization problems that has proved to be profitable to treat classical Traveling Salesman Problems.^{3,4} It is important to remark that for various NP-hard problems Rajasekaran et al.⁵ presented empirical outcome obtained by means of SA, where they reported to have achieved better results than those yielded by other global-optimization heuristics.

In 1983 Kirkpatrick et al.³ stated that the detailed analogy with annealing in solids provided by the SA methodology is a useful framework for the optimization of the properties of very large and complex systems. SA gives the great advantage of being easily applicable to problems with a combinatorial structure because SA is simple to be implemented. Besides, it is possible to combine SA with other heuristic techniques, such as expert systems.⁶ From then on, various kinds of problems⁶ have been addressed and solved successfully with SA. Aarts et al.⁷ wrote a review and some problems related to the design of transport networks are remarkable.^{8–10} Eglese et al.¹¹ described thoroughly the “pure” SA algorithm and made some suggestions on how to improve its functioning. Bertsimas et al.¹² analyzed SA's convergence capacity and discussed SA's behavior in practice.

In this work we have taken advantage of SA's inherent qualities in order to efficiently solve an NP-hard optimization problem that consists of finding the most convenient location for a pipeline network. In particular, the case under detailed analysis is about hydrocarbon pipelines that would be spread out onshore in Argentina where there are 19 sedimentary basins with a total surface of about 1 750 000 km². Nowadays, only the following five basins (called Cuencas in Spanish) are hydrocarbon productive: Noroeste, Cuyana, Neuquina, Golfo de San Jorge, and Austral. These basins have a lot of productive wells that are widespread and far from the processing plants or consumption centers. Thus, it is essential to optimize the transport lines and the location of the stores and the process plants in order to maximize this business' profitability. Cuenca Austral, which stretches from the South of Santa Cruz province

Received: January 3, 2013

Revised: April 20, 2013

Accepted: May 24, 2013

Published: May 24, 2013

into the North of Neuquén, is a rich basin that could hold over $401\,308 \times 10^3 \text{ m}^3$ of oil and $358\,726 \times 10^6 \text{ m}^3$ of gas reserves. Natural gas is mainly composed of methane (CH_4 , 80–95%), ethane (C_2H_6 , 4–8%), higher hydrocarbons, and other minor compounds, such as N_2 , CO_2 , and H_2S . Once extracted from the wells, it has to be treated in various conditioning and separation stages so as to obtain different fractions. There are basically two separation processes: Primary and Secondary Separation. The liquid fraction obtained in the Primary Separation is called condensate, which is mainly composed of propane, butane, and higher hydrocarbons. Later, this condensate is subjected to a stabilization stage, where some natural gasoline, which is composed of pentane and higher hydrocarbons, is obtained. The gaseous product from the Primary Separation step still contains liquefiable components. Then, it is sent to a Secondary Separation stage, where three fractions are separated: ethane, propane + butane, and natural gasoline. In general, between this pair of separation processes, there are various conditioning steps.¹³

In Argentina natural gasoline production amounted to 1 782 812 m^3 during 2010.¹⁴ This natural gasoline might be employed as a raw material in order to add value to the final products. A preliminary economic analysis¹⁵ showed that it is a profitable alternative to use it for olefin production (ethylene) by thermal cracking. Then, it is fundamental to define its transport routes. Our case study takes place in the Province of Santa Cruz; where there are two basins producing oil and gas: Cuenca Austral and Cuenca Golfo San Jorge. In 2010, approximately 30% of the total natural gasoline was extracted from those places.¹⁴ This amount makes these basins the second largest natural gasoline source in Argentina, just after the fields in Cuenca Neuquina.

In Santa Cruz there are no specific transport pipelines installed for natural gasoline, no process plants, and no shipment places assigned. Nowadays, natural gasoline is carried by truck to be mixed with crude oil. Therefore, the installation of specific pipeline networks will also provide safer operational conditions.

The logistic problem addressed here is very important not only for Argentina, but also for other countries that hold widespread sources to exploit. By way of illustration, oil shales are widely distributed in the United States and may be a future source of gas and petroleum (e.g., Marcellus Shale, Haynesville Shale, Bakken Shale, etc.).

Over the last decades several transport problems have been solved by means of various optimization techniques. In general planning and scheduling have been addressed globally. Then, these tasks have included the determination of market needs, the compliance with existing regulations, environmental risks, inventory analysis, the specification of pipes and components, pipeline design, and route selection.

In some cases, the pipelines are already installed and their most profitable use should be determined. For example, Boschetto et al.¹⁶ have proposed an operational generic mixed-integer linear planning model that optimizes the volume of products and the flow directions in existing pipelines so as to attain storage goals, while taking into account the consumers' demands. In contrast, Cafaro et al.¹⁷ have presented a novel formulation about the short-term operational planning of a tree-structure pipeline so as to find the optimal schedule of pumping and delivery operations that minimizes operating costs.

In particular, pipeline network design has been included in some optimization models for planning and scheduling. Moura

et al.¹⁸ proposed a novel combinatorial hybrid algorithm that combines a constructive heuristic strategy with a constraint programming model to determine the best way to transport different products in a pipeline network in order to supply market demands, while also satisfying hard operational constraints related to product sequencing, flow rates, and tank capacities. Van der Heever et al.¹⁹ determined the system connectivity via pipelines by using a multiperiod mixed-integer nonlinear programming model for the long-term design and planning of offshore hydrocarbon field infrastructures with complex economic objectives. In contrast, Iyer et al.²⁰ also contemplated pipeline infrastructure for a given planning horizon and they obtained satisfactory results for a simpler economic model, resorting to a decomposition strategy.

Manenti et al.²¹ highlight the relevance of the strategic coordination of enterprise-wide decentralized productions, pointing out that, for supply chain optimization, the interactions between suppliers and customers are incorporated into a well-defined decision making process. Moreover, Seuring²² enhances the significance of integrated supply chain management (SCM). An integrated SCM considers industrial ecology, also capturing strategic and tactical decisions, such as geographic distribution, hedging via future contracts, routing to plant sites, and warehouse positioning.²³

A network design problem concerns topology determination, i.e. basically the configuration of the links between nodes. This is closely related to several factors, such as expected applications, node distances, physical connections, transmission rates, and fluid properties. Naturally, there are multiple variants in the literature about optimal design of a given network. For example, Andre et al.²⁴ considered in a single optimization model the topology and the pipeline dimensioning. De la Cruz et al.²⁵ have only addressed topological optimizations for pipeline networking, while Adb-El-Barr²⁶ has covered another field of application: the classical communication networks. Moreover, Gonzalez et al.²⁷ have considered more details in their model, including the dynamic characteristics for a gas distribution pipeline network.

As to the algorithmic approach, Wei et al.²⁸ presented a hybrid Genetic-Simulated Annealing algorithm. Nevertheless, individual algorithms have also shown a satisfactory performance. De la Cruz et al.²⁵ defined genetic operators accordingly and found the distribution of an oil-pipeline network by means of their evolutionary algorithm. In turn, Marcoulaki et al.²⁹ employed SA in an optimization framework for the routing and equipment design of main pipelines.

In this work a tactical approach based on the hybridization of the Simulated Annealing³⁰ metaheuristic together with the General Algebraic Modeling System (GAMS)³¹ is adopted in order to solve a pipeline network design problem. The problem formulation is detailed in Section 2. The strategy is outlined in Section 3. Then, some relevant details about our tactics to handle geographic information are considered. Our procedure is discussed in Section 5, where it is illustrated by means of a couple of cases. Lastly, the computational performance is analyzed in Section 6. The last section refers to the main conclusions.

2. PROBLEM STATEMENT

In this paper, we have considered the design of pipeline networks for hydrocarbons. We have posed this issue as a main location problem combined with transport features. Our objective is to minimize the installation costs (Z) so as to

efficiently send the production from the oil fields (S_i) to the prospective stores or processing plants (P_j). The oil fields are modeled as nodes, where many closely located wells may be gathered to constitute a single node. In fact, these nodes are anchored. In contrast, the model itself evaluates the location of concentrating nodes (intermediate stops) that collect the flow to have it then sent somewhere else. This combinatorial problem consists of the determination of a subset of multiple options for the possible places whose cost varies according to their geographic location.

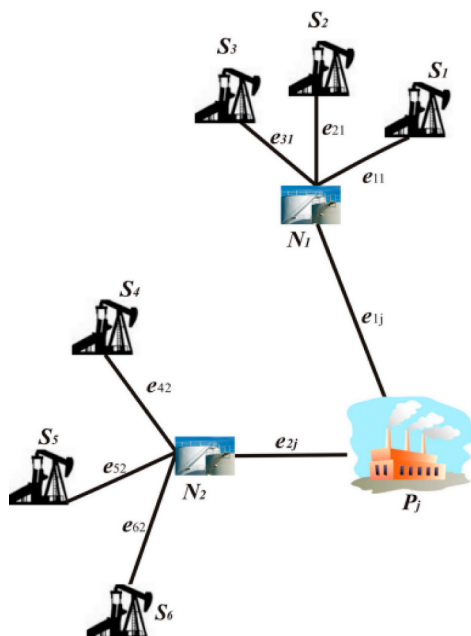
The need for the incorporation of concentrating nodes where the hydrocarbons extracted from the fields might be stored to be eventually transported to a processing plant triggers a location problem. It has been assumed that all the standard concentrating nodes receive the same hydrocarbon volume. The installation cost β of a concentrating node was calculated with the classical method³² for the variation of equipment cost with size. For network design purposes, the product accumulation in the concentrating nodes was not taken into account. Nevertheless, the storage tanks were considered in β .

Besides, for each potential solution, i.e. for each subset of concentrating nodes, a transport plan becomes necessary so as to minimize the total cost of both the proposed installations and the production transport. Then, once the amount of concentrating nodes together with their locations has been decided, a transport problem should be solved in order to provide an optimal product distribution. In fact, the existence of intermediate stores prompts speculation over transshipment problems. Nevertheless, any transshipment problem may be easily converted into an equivalent transport problem.³³ It should be noted that an equivalent problem yields the same solution as the original one.

The problem formulation is based on the following assumptions:

- A graph representation is employed in order to map the distribution (see Scheme 1)

Scheme 1. Graph Representation for a Typical Oil Field Infrastructure



- This notation was adopted for the paths: the edges e_{ij} run from node i to node j .
- Any oil field may act as a concentrating node.
- As to the concentrating nodes (intermediate collectors), their input flow rate is equal to the output flow rate. In this way, all the hydrocarbon's production is sent somewhere, i.e. either to another node or a plant.
- The plants contemplate no hydrocarbon's output. They only receive all the product, in order to have it processed.
- The cost C_{ij} was calculated as the pipeline construction cost running along the e_{ij} . These costs are summarized in a symmetric matrix C , whose elements are C_{ij} .
- There may be an interconnection among the concentrating nodes or among the oil fields.
- Each concentrating node is a binary variable called N_k , where $N_k = 1$ when the k th node is being employed and $N_k = 0$ when it is not.
- There is a weight ω_k associated with the k th node that is calculated on the basis of a difference in heights (eq 1). The heights above sea level H_k were obtained through gvSIG.³⁴

$$\omega_k = \frac{H_k}{H_{\max}} \quad (1)$$

- For liquids, the weights ω_k take into account the pumping difficulty. For gases they are equal to 1.
- The pipeline size was considered a parameter in our optimization problem. For the network under study, the pipeline diameters were calculated from the fluid speed and the production of natural gasoline. In Argentina, the node interconnections are generally installed by using pipelines of diameter either 4 in. or 8 in. Nevertheless, the algorithm presented in this paper can be applied to any diameter because the diameters are input data.

The most important decision variables are the concentrating nodes (N_k) and the amount of transported hydrocarbons from i to j (X_{ij}). These amounts are continuous variables between some minimum (X_{ij}^{\min}) and maximum (X_{ij}^{\max}) positive values.

The mathematical model aims at minimizing costs (eq 2) to transport the products from the oil fields to the stores. The parameter m is the addition of the numbers of oil fields, concentrating nodes, and plants.

$$\min Z = \sum_{i=1}^{f+n} \sum_{j=1}^m C_{ij} X_{ij} + \beta \sum_{k=1}^n \omega_k N_k \quad i \neq j \quad (2)$$

where $m = f + n + p$

The first constraints (eq 3) ensure that the total amount of transported product is equal to the offer (A_i). The second set of constraints (eq 4) restricts the fluid quantity to the stated amount (B_j). Finally, there are constraints (eq 5) to ensure that the decision variables (X_{ij}) are non-negative.

$$\sum_{j=1}^m X_{ij} = A_i \quad 1 \leq i \leq f + n \quad (3)$$

$$\sum_{i=1}^{f+n} X_{ij} \leq B_j \quad f + n + 1 \leq j \leq m \quad (4)$$

$$X_{ij} \geq 0 \quad (5)$$

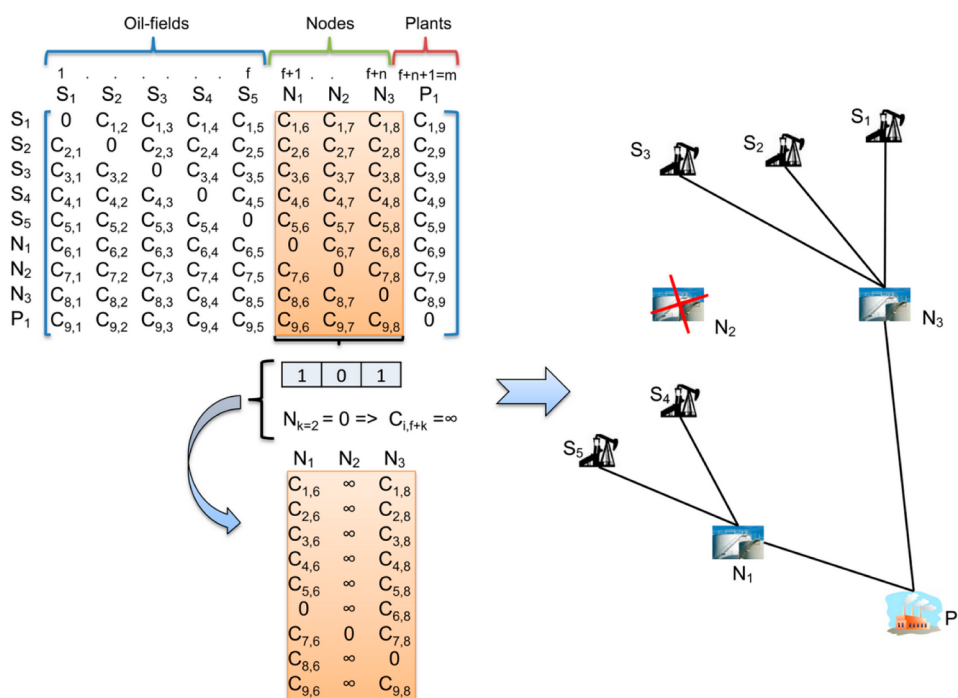
Scheme 2. Pseudoalgorithm for the SAG

Procedure Simulated Annealing (T_0 , Iter, Cad)

```

Temp =  $T_0$                                      % A temperature is initialized
Xactual = Initialization ()                     % A random solution is obtained
for i = 1 to Iter do                             % The number of iterations is established
  for j = 1 to Cad to                             % The temperature is kept
    Xnew = Perturb(Xactual)                       % A neighboring solution is generated
    Delta = f(Xnew) - f(Xactual)                 % The difference between solutions is evaluated
    if (  $U(0,1) < e^{(-Delta/Temp)}$  ) or (Delta < 0) then % The acceptance criterion is applied
      Xactual = Xnew                             % The solution is updated
    end if
  end for
  Temp =  $T_0 / (i + 1)$                          % The cooling temperature is decreased
end for

```

Scheme 3. A Small Pipeline Network: Management of the Cost-Matrix Change When Node N_2 is Deactivated

This model can be applied to liquid or gas hydrocarbons at any locations. These features only affect the costs C_{ij} , which are calculated independently from the model. The installation of concentrating nodes and the pipeline network may be directly affected by other factors, such as political decisions, environmental regulations, and rights of way through private property. In this paper these factors were not taken into account.

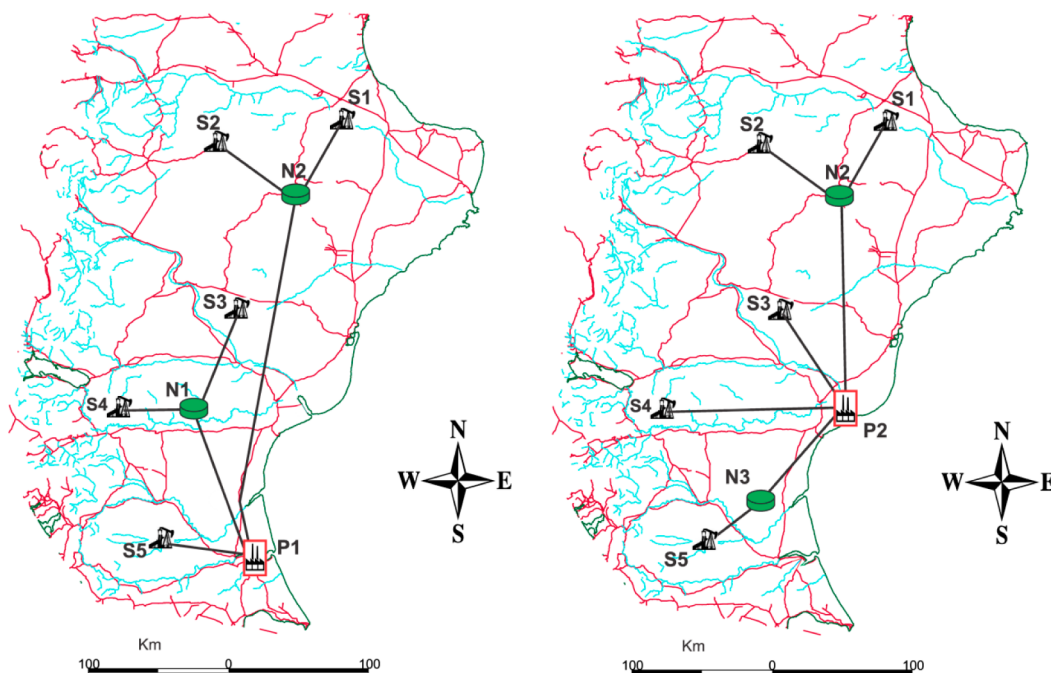
This is formulated as a combinatorial optimization problem. The resulting system is solved by means of a metaheuristic technique of Simulated Annealing with GAMS (SAG). It has two main stages: a Simulated Annealing procedure for site location and a linear programming cost minimization procedure (called TRANBOR1) that solves the transport problem for each chosen location employing tools provided by GAMS software.

3. COMPUTATIONAL STRATEGY

When it is desired to solve an NP-hard problem of large scale, there are two suitable options: either (1) to choose an

algorithm that ensures getting an optimum in an excessively long computational time, or (2) to employ a technique that yields a feasible solution in a polynomial computer time, though finding the global optimum value is not guaranteed.

In this work, the formulated problem has been considered as NP-hard.³⁵ Then, getting feasible solutions through heuristic methods is plainly justified in view of the problem's complexity. The definition of problem complexity³⁶ refers to the amount of resources necessary to carry out a calculation. When solving a problem, the resources commonly addressed are the time (number of algorithmic execution steps) and the space (required amount of memory). It is widely believed that there is no algorithm that allows determining the optimal solution of a NP-hard problem in a polynomial time.^{37,38} This means that the process of finding the solution grows exponentially with respect to problem's size. Then, it implies a significant amount of computational time in the algorithm's execution for large problem instances, which is a hindrance that has led researchers to consider being satisfied with suboptimal

Scheme 4. Optimal Pipeline Routes for Example 1^a

^aCase I (on the left): Plant located in Rio Gallegos. Case II (on the right): Plant located in Puerto Santa Cruz.

solutions obtained by heuristic methods or approximate methods that make use of heuristics.

Since its appearance in 1983, SA has been successfully applied to a wide variety of combinatorial optimization problems in various fields.^{39,40} It can be stated that SA is one of the most applicable and flexible algorithms. Though optimality cannot be ensured, SA quickly provides solutions that behave favorably. In particular, Rodríguez et al.⁴¹ presented a Vehicle Routing Problem (VRP) with a big set of possible solutions. The VRP required a technique that efficiently explored the search space. SA competed with a genetic algorithm and SA succeeded in yielding better solutions in a shorter computational time.

The devised metaheuristic algorithm is essentially a hybridization of Simulated Annealing with the GAMS package. SA is employed for the allocation of hub nodes, whereas GAMS is used to solve the transportation problem that provides optimal product distribution for a given configuration.

The SAG algorithm finds the adequate amount of concentrating nodes and their most convenient location. The space of solutions is represented by means of binary vectors v . Each vector v of n bits corresponds to n feasible locations for the concentrating nodes, where an element is assigned as the number one to each particular node whose installation has been proposed. SAG's procedure is summarized in the pseudoalgorithm given in Scheme 2.

GAMS allows solving problems in a quick and efficient way. Its main advantage is that it provides the distinct possibility of applying different solvers in the generated models, e.g. linear, nonlinear, integer models, etc., along with the allowance to switch between these possibilities by making small changes to the input file. To model the case studies in GAMS, we have sought to simplify the analysis. The same output was considered for all fields, carrying all the production to the destination plant. The construction cost of each concentrating node was calculated based on the height above sea level of the

zone where it is located. Finally, we have taken into account the distances and height differences in order to calculate the cost of carrying the production from one point to another one.

For every potential solution SA generates a modification of cost matrix C . Subsequently, GAMS employs the modified C matrix in order to evaluate the corresponding fitness value Z . When the k th concentrating node is not chosen ($N_k = 0$), all the entries C_{ij} with $i = 1, \dots, m$ and $j = f + k$ are penalized with an extremely high value so that N_k is never taken into account. Scheme 3 represents a small example where the algorithmic management of the node configuration via C is illustrated. When a concentrating node is not considered in the configuration, e.g. N_2 , the column that corresponds to that node assumes extremely high values, i.e. the entries C_{i7} ($i = 1, 9$) tend to infinity. In this way, when GAMS employs the modified C matrix, N_2 will not be chosen because it is costly. Scheme 3 also shows the resulting network for $N_2 = 0$.

In short, the associated cost matrix C is generated only once, at the beginning of the procedure. Then, for each stage, its elements are modified according to the active nodes. The technique saves storage and avoids additional data structures that would affect computational performance.

4. GEOGRAPHICAL FEATURES

The geographical features were introduced in order to take into account the landscape involved within the area. The gvSIG tool³⁴ was employed to capture the site complexity. This is a free open-access geographic information system (GIS). This work was specifically carried out for onshore facilities because the considered shapes were taken with terrestrial topographical charts. Building costs significantly increase when a pipeline has to cross a river or a road. Moreover, the difference in heights was considered because it influences the pumping costs of natural gasoline.

In the analysis exemplified in this paper the information about the geographical features that characterize the province of

Santa Cruz (rivers, routes, oil fields, etc.) was taken from the available databases^{42–44} on the Internet. We defined a map with these data by using gvSIG. This facility let us build the corresponding distance matrix **D**. Besides, the pipeline diameters (IDs) were estimated from the historical production of each oil field. Finally, the cost matrix **C** was calculated⁴⁵ on the basis of **D** and IDs.

In the generation of the values assigned to **C**, the following requirements were also considered:

- It is convenient to favor the installation of concentrating nodes, strategically planning that the raw material might be stored there.
- When the plant is located close to a populated area or any port, it is unadvisable to send too many pipes to the plant so as to minimize the environmental impact.

Besides, it is important to point out that the installation of concentrating nodes is central to improve plant flexibility. For operational purposes, they serve to build up reserves. Then, this extra storage may ensure productivity. For instance, if a well stops its production for any reason, the plant operation will remain unaffected by that event. Moreover, the concentrating nodes are useful for safety purposes because they allow shifting product accumulation toward less risky zones, i.e. away from petrochemical plants or crowded areas.

5. CASE STUDIES

In this section, the performance of the proposed formulation is evaluated by solving three examples all dealing with the tactical design of pipeline networks. The case studies are modified instances of a real-world problem under consideration. It is the design of a pipeline network that would carry natural gasoline from the fields to a potential processing plant. This plant might be located in either Puerto Santa Cruz (P_2) or Río Gallegos (P_1). Both locations were chosen for this analysis because they have proved to be favorable.⁴⁶ In Puerto Santa Cruz there are industrial projects that would make exports viable thanks to the industrial growth in the area. In turn, in Río Gallegos there is a port that nowadays exports primary products coming from the oil fields.

Example 1 involves five theoretical oil fields, three concentrating nodes, and two potential plant locations. The problem under consideration is the design of a pipeline network that will carry natural gasoline from these fields to the processing plants. For each of the prospective locations, the optimal alternative routes found with the proposed methodology are depicted in Scheme 4.

The computational testing is summarized in Table 1, where the total cost Z (eq 2) corresponds to the fitness values. Z takes

Table 1. Fitness Values for Example 1

case	N_1	N_2	N_3	fitness (Z) (10^6 \$)	distance (km)
I	1	1	0	463	1176
II	0	1	1	375	1055

into account the pipeline-construction costs and the installation cost of each concentrating node. For the sake of comparison, several runs were carried out for the processing plant located in either Río Gallegos (case I) or Puerto Santa Cruz (case II). Regarding pipeline-installation costs, Scheme 4 shows the cheapest configurations in both cases. The values can be compared from Table 1, where it can be noticed that case II is

more convenient because the node configuration makes the pipeline length and the installation costs of the individual nodes relatively inexpensive. Moreover, it should be noted that in this region the runs indicated that it is worthwhile installing two concentrating nodes. At first sight, for case II the presence of a concentrating node in the middle of the pipeline linking S_5 with P_2 is curious. Nevertheless, building a long pipe that links S_5 with plant P_2 without using a concentrating node in the middle proved to be more expensive because the costs were also affected by the requirements mentioned at the end of Section 4.

Let us turn to more complex realistic cases. Example 2 addresses the same region, but the existing oil fields that are active at present are considered. Besides, the problem size (21 existing fields) is greater than that in Example 1 (5 theoretical fields). Table 2 shows the best four alternatives found for

Table 2. Fitness Values and Distance Covered by the Optimal Networks Calculated for Example 2

alternative	concentrating nodes	fitness (10^6 \$)	distance (km)
A	N_6, N_{24}	623	1796
B	N_6, N_{26}	630	1645
C	N_5, N_{24}	644	1680
D	N_5, N_{26}	651	1529

Example 2. Options A and C profit from the low installation cost of concentrating node N_{24} , while options B and D benefit from the shorter path required to mount all the network. In contrast with the total cost of the network, node building is relatively expensive. Therefore, the installation of several nodes is not promoted. Nevertheless, there are two remarkable zones that concentrate all the wells due to the distribution of the existing fields. Then, the algorithm logically distinguishes the installation of a concentrating node in each of those zones as the best economic alternative (see Scheme 5). The most convenient configuration is Alternative A due to its low fitness value. Therefore, nodes N_6 and N_{24} are adopted. From an economic point of view, it is convenient to send all the production to a single processing plant. So, the next step is to determine the most suitable ending path.

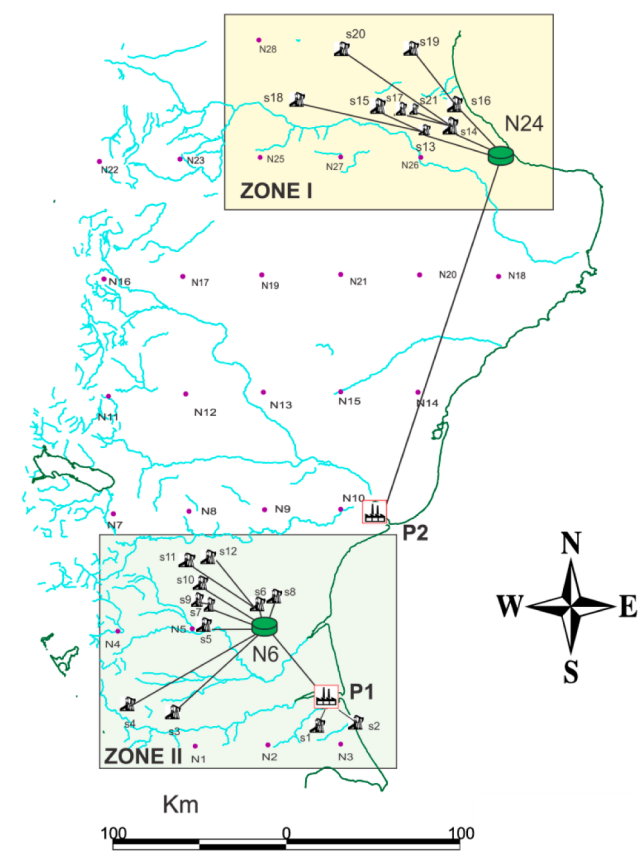
As to the transport toward the final processing plant (Scheme 6), Table 3 shows the best alternatives for each case. Both make use of the same pair of concentrating nodes (N_6, N_{24}), but there is a difference between the total cost of the networks. Then, Case I is preferable.

Compared with Example 2, a major change has been introduced in Example 3. A potential exploitable well (S_{22}) has been incorporated so as to consider future drillings (Scheme 7) that are likely to be necessary when initial stocks at existing wells become scarce to meet consumers' demands. The zone selected for the location of this prospective well was determined within the exploration areas for new sources of hydrocarbons obtained from online geographic information.⁴³ Table 4 shows the results for the processing plant P_1 . As expected, the total installation cost for this network is higher, but this time the best choice is to make use of N_5 . It is clear that the network design depends on future developments. Therefore, the decisions should be made according to strategic planning.

6. COMPUTATIONAL PERFORMANCE

Regarding validation, Examples 1, 2, and 3 may be considered as small toy examples. A comparison is presented in Table 5

Scheme 5. Distribution Network Proposed As Alternative A for Example 2



between their optimal solution obtained by using brute force and the results yielded by our model. For each example

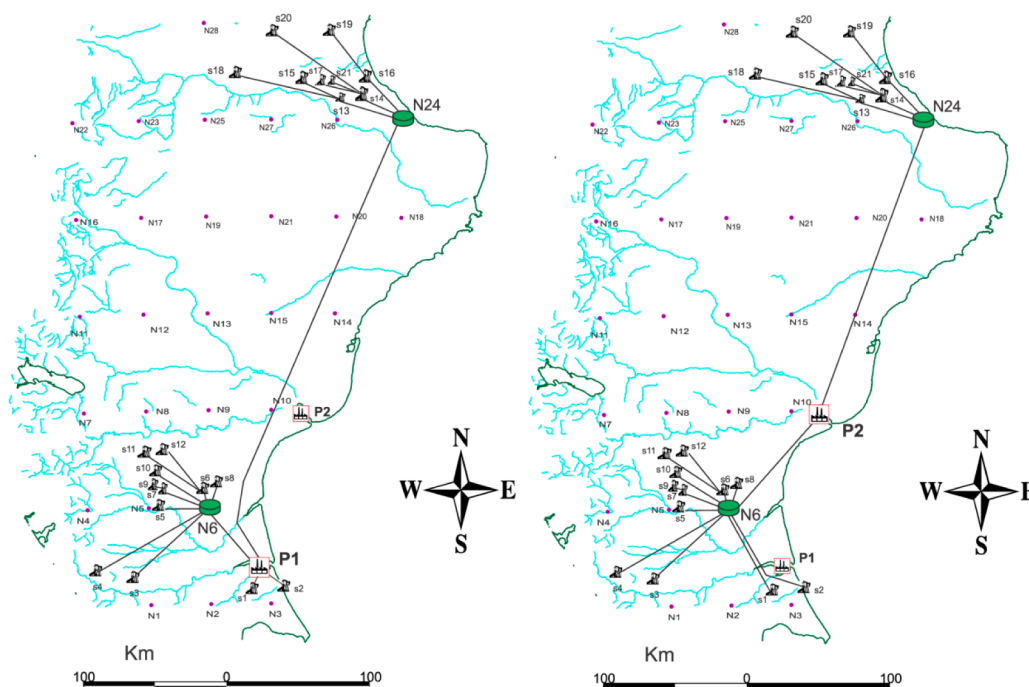
Table 3. Cost and Distance Covered by the Optimal Networks for Example 2 with Nodes in N_6 and N_{24}

case	destiny	fitness (10^6 \$)	distance (km)
I	Río Gallegos (P_1)	639	1981
II	Puerto Santa Cruz (P_2)	674	2030

described in Section 5 thirty executions were carried out. The exact solution coincided with the best solution found via SAG. Our proposed SAG metaheuristics was implemented in JAVA and tried out in a PC with 1 GB RAM and a Core 2 Duo 2.53 GHz processor. After relatively short computing times, the results yielded by the SAG algorithm were accurate enough.

Table 5 provides evidence that SAG performed well in obtaining good solutions within reasonable computing times. For comparison purposes, the optimal solution is shown in the fifth column. The values reported in the sixth column are equal to the average fitness value of 30 executions. SAG proved to be robust because, for all these tests, the optimum solution was obtained on more than 60% times, reaching the rest of the times values that were very close to the optimum. Then, by repeatedly executing SAG, the solution that appears most often can be adopted as the best one. Example 1 is the smallest case and the media coincided with the best solution. When the size increased, there was a difference between the best solution and the media. This is to be expected because the amount of combinations augments exponentially with respect to the input size.

As to the optimal solution for example 3, it was obtained by trying out all possible combinations in GAMS. In this exhaustive search, the required computational time was about 61 200 s. In comparison, this amount is about 76 500% higher than the computational time spent by the SAG algorithm for this small example. Then, SAG proved to be an efficient procedure. Naturally, when the problem size is even greater, it

Scheme 6. Optimal Pipeline Routes for Example 2^a

^aCase I (on the left): Plant located in Río Gallegos. Case II (on the right): Plant located in Puerto Santa Cruz.

Scheme 7. Optimal Pipeline Routes for Example 3

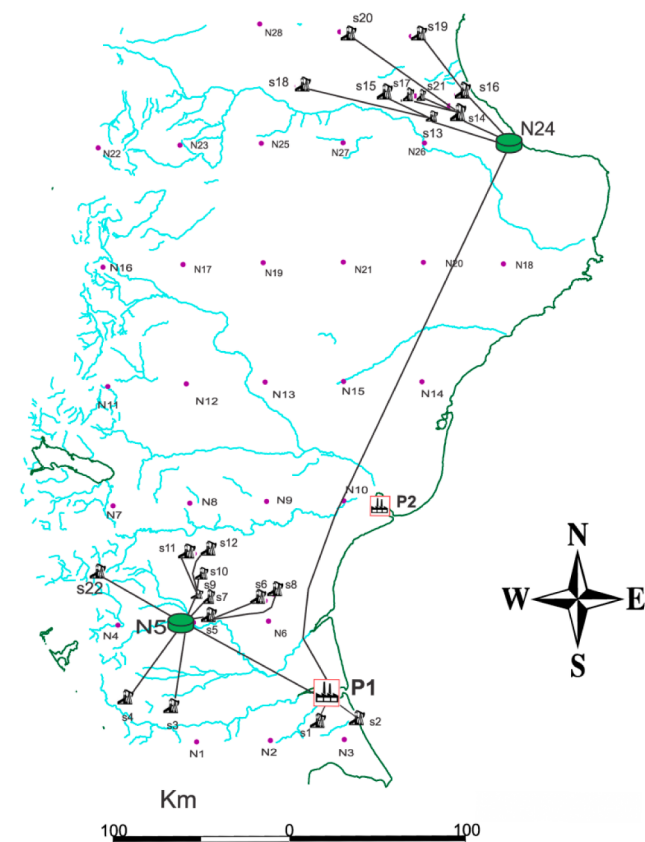


Table 4. Cost and Distance Covered by the Optimal Networks for Example 3

concentrating nodes	fitness (10^6 \$)	distance (km)
N ₅ , N ₂₄	673	1953
N ₆ , N ₂₄	677	2065

Table 5. Statistics for the Examples 1, 2, and 3 Solved via SAG

example	case	oil fields	concentrating nodes	optimal solution (10^6 \$)	media (10^6 \$)	CPU time (s)
1	II	5	3	360	360	1.27
2	A/I	21	28	623	634	73.28
3	I	22	28	660	665	80.65

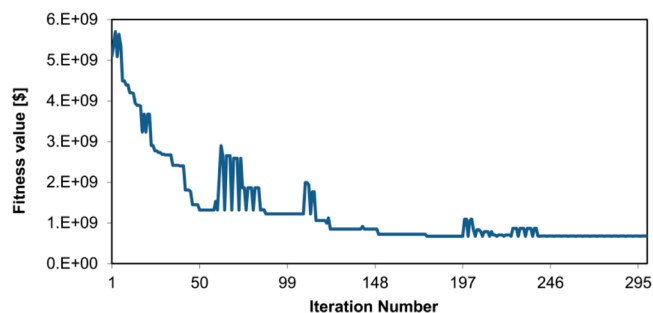
is prohibitive to attempt to find the exact solution by examining all possible combinations. Consequently, the employment of an efficient search procedure, such as the one proposed in this paper, is advisable.

The metaheuristics' profitability was directly affected by the values chosen for the parameters. These values were fixed after a preliminary testing phase, where we studied the quality of the obtained solutions and the computational time that was required to get them. According to this search, the best values adopted for these examples are the following:

- temperature descent = 0.99
- length of Markov's chain = 3; (tested range: [1, 10])
- iteration number = 100; (allowed range: [50, 300])

Scheme 8 shows the algorithmic behavior in a single execution. The fitness value decreases sharply during the first

Scheme 8. Convergence Behavior of SAG Algorithm



half of the iterations, while there is a peak whenever a worse solution is accepted on the lookout for a new search path. Besides, the exploitation of the most useful areas is reinforced in the second half of the iterations. Hence, the probability of accepting a worse solution is low. In this way, the algorithm is flexible and frequently succeeds in escaping from local optima, thus improving its performance considerably.

6. CONCLUSIONS

In this paper, a general metaheuristic algorithm based on the simulated annealing approach, which was called SAG, is proposed for the general optimization of the distribution of pipeline networks. The design also takes into account the geographic features of the zone under study. The goal is to minimize the total pipeline-routing cost, at the same time pinpointing the strategic location of concentrating nodes. As to optimization solving, this simulated annealing approach is a promising layout for time improvements.

By way of illustration about several attributes of this proposal, a representative variety of cases were tackled. All of them could be solved successfully in a reasonable amount of CPU time. Among the general-model features highlighted by the examples, the following remarkable milestones especially deserve mentioning:

- an automatic analysis of the addition of intermediate nodes when initial stocks at existing wells are scarce to meet consumers' demands
- an effective management of widespread sources
- cost savings in pipeline installation
- a novel hybrid approach, with a noncomplex model formulation that serves to reach the desired target with flexibility.

SAG's robustness was shown by the analysis of extra results. SAG is robust because a high percentage of the executions could find the optimum, or else solutions very close to the optimal values. Besides, SAG was carefully programmed to ensure efficiency and flexibility. SAG is efficient because it required comparatively short computational times, while it is flexible for it succeeded in escaping from local optima until a good solution was found.

AUTHOR INFORMATION

Corresponding Author

*E-mail: poteiza@plapiqui.edu.ar. Tel:054-0291-4861700, Int: 250.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We gratefully acknowledge Surveyor Luis Ariel Raniolo, who is a Professional from the Argentinian Oceanography Institute (IADO), CCT-CONICET – Bahía Blanca. Besides, we thank Mg. Carlos Eugenio Gigola (PLAPIQUI) and Mg. Juan Ignacio Ardenghi (UNS) for their help at covering fundamental material.

NOMENCLATURE

- A_i : natural gasoline offer from the i th oil field (t)
 B_j : natural gasolines' demand of the j th processing plant (t)
 β : installation cost of a standard concentrating node (\$)
 C_{ij} : transport cost from the i th node to the j th node (\$/t)
 C : cost matrix with C_{ij} entries, $C \in \mathcal{R}^{m \times m}$
 f : number of oil fields
 D : distance matrix
 H_k : the height of the k th concentrating node
 H_{\max} : the highest node of all the concentrating nodes
 m : the addition of the numbers of oil fields, concentrating nodes, and plants (i.e., $m = f + n + p$)
 n : number of concentrating nodes
 N_k : the k th concentrating node
 p : number of plants
 S_i : the i th oil field
 X_{ij} : amount of transported gasoline from the i th node to the j th node (t)
 ω_k : weight for the k th concentrating node
 Z : fitness value representing the total cost (\$)

REFERENCES

- (1) Hajek, B. Cooling schedules for optimal annealing. *Math. Oper. Res.* **1988**, *13* (2), 311.
- (2) Blondel, V.; Tsitsiklis, J. N. NP-Hardness of some linear control design problems. *SIAM J. Control Optim.* **1997**, *35*, 2118.
- (3) Kirkpatrick, S. C. D.; Gelatt, Jr.; Vecchi, M. P. Optimization by simulated annealing. *Sci.* **220** **1983**, 671.
- (4) Cerny, V. A thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm. *J. Opt. Theor. Appl.* **1985**, *45* (1), 41.
- (5) Rajasekaran, S.; Reif, J. H. Nested annealing: A provable improvement to simulated annealing. *Theor. Comput. Sci.* **1992**, *99* (1), 157.
- (6) Elmohamed, M. A. S.; Coddington, P.; Fox, G. A comparison of annealing techniques for academic course scheduling. *Lect. Notes Comput. Sci.* **1998**, *1408*, 92.
- (7) Aarts, E. H. L.; Korst, J. H. M. Selected topics in simulated annealing. In *Essays and Surveys in Metaheuristics Series*; Kluwer Academic Publishers: Boston, MA, 2002.
- (8) Costa, A. L. H.; Medeiros, J. L. D.; Pessoa, F. L. P. Optimization of pipe networks including pumps by simulated annealing. *Braz. J. Chem. Eng.* **2000**, *17* (4–7), 887.
- (9) Yaghini, M.; Momeni, M.; Sarmadi, M. A simplex-based simulated annealing algorithm for node-arc capacitated multi-commodity network design. *Appl. Soft Comp.* **2012**, *12* (9), 2997.
- (10) Rui, X.; Gongyuan, L.; Dong, H. Simulated annealing algorithm for solving a bi-level optimization model on high-speed railway station location. In *Third International Conference on Information and Computing*, Wuxi, China, June 4–6; Jiangnan University: Wuxi, China, 2010; pp 159–162.
- (11) Eglese, R. W. Simulated annealing: A tool for operational research. *Eur. J. Oper. Res.* **1990**, *46*, 271.
- (12) Bertsimas, D.; Tsitsiklis, J. Simulated annealing. *Stat. Sci.* **1993**, *8* (1), 10.
- (13) Kidnay, A. J.; Parrish, W. *Fundamentals of Natural Gas Processing*, 1st ed.; CRC Press Taylor & Francis Group: London, 2006.
- (14) IPA. *Información Estadística de la Industria Química y Petroquímica de la Argentina*, 31st ed.; 2011.
- (15) Cañete, B.; Oteiza, P. P.; Gigola, C. E.; Brignole, N. B. Gasolina natural: Un sustituto atractivo para la producción de etileno en Argentina. *Rev. Petroq. Petról. Gas y Qca.* Summary in Technical Scientific Work Section, **2012**, 283, 206. Full article available at <http://revistapetroquimica.com/gasolina-natural-un-sustituto-atractivo-para-la-produccion-de-etileno-en-argentina/>, 2013.
- (16) Boschetto, S. N.; Magatão, L.; Neves-Jr, F.; Barbosa-Póvoa, A. P. F. D. MILP planning model for real-world multiproduct pipeline network. In *21st European Symposium on Computer Aided Process Engineering*, Porto Carras Resort, Chalkidiki, Greece, May 29–June 1, 2011; Vol. 29, pp 995–999.
- (17) Cafaro, D. C.; Cerdá, J. A rigorous mathematical formulation for the scheduling of tree-structure pipeline networks. *Ind. Eng. Chem. Res.* **2011**, *50*, S064.
- (18) Moura, A. V.; de Souza, C. C.; Cire, A. A.; Lopes, T. M. T. Heuristics and constraint programming hybridizations for a real pipeline planning and scheduling problem. In *11th IEEE International Conference on Computational Science and Engineering*, Campinas, Brasil, July 16–18, 2008; pp 455–462.
- (19) Van der Heever, S.; Grossman, I.; Vasantharajan, S.; Edwards, K. A. Lagrangean decomposition heuristic for the design and planning of offshore hydrocarbon field infrastructures with complex economic objectives. *Ind. Eng. Chem. Res.* **2001**, *40*, 2857.
- (20) Iyer, R. R.; Grossmann, I. E. Optimal planning and scheduling of offshore oil-field infrastructure investment and operations. *Ind. Eng. Chem. Res.* **1998**, *37*, 1380.
- (21) Manenti, F.; Bozzano, G.; D'Isanto, M.; Nascimento, L. N. M.; Zuniga, L. L. Raising the decision-making level to improve the enterprise-wide production flexibility. *AIChE J.* **2012**, *1*.
- (22) Seuring, S. Integrated chain management and supply chain management comparative analysis and illustrative cases. *J. Clean. Prod.* **2004**, *12*, 1059.
- (23) Varma, V. A.; Reklaitis, G. V.; Blau, G. E.; Pekny, J. F. Enterprise-wide modeling & optimization - An overview of emerging research challenges and opportunities. *Comput. Chem. Eng.* **2007**, *31* (5–6), 692.
- (24) Andre, J.; Auray, S.; Brac, J.; Wolf, D. D.; Maisonnier, G.; Sidi, M. O.; Simonnet, A. Design and dimensioning of hydrogen transmission pipeline networks. *Eur. J. Oper. Res.* **2013**, *229*, 239–251, DOI: <http://dx.doi.org/10.1016/j.ejor.2013.02.036>.
- (25) De la Cruz, J. M.; de Andres-Toro, B.; Herrán, A.; Besada Porta, E.; Fernandez Blanco P. Multiobjective optimization of the transport in oil pipelines networks. *Emerging Technologies and Factory Automation, IEEE Conference*, Lisbon, Portugal, 16–19 September, 2003; Calouste Gulbenkian Foundation: Lisbon, Portugal, 2003; vol. 1, pp 566–573.
- (26) Adb-El-Barr, M. Topological network design: A survey. *J. Net. Comput. Appl.* **2009**, *32*, S01.
- (27) Herrán-Gonzalez, A.; De La Cruz, J. M.; De Andrés-Toro, B.; Risco-Martín, J. L. Modeling and simulation of a gas distribution pipeline network. *Appl. Math. Model.* **2009**, *33*, 1584.
- (28) Wei, L.; Jiang, H.; Liu, Y. Hybrid genetic-simulated annealing algorithm of location-allocation, optimization of looped gathering and transportation pipe network. In *Fifth International Conference on Natural Computation*, Tianjian, China, 14–16 August, 2009; vol. 114, pp 275–280.
- (29) Marcoulaki, E. C.; Papazoglou, I. A.; Pixopoulou, N. Integrated framework for the design of pipeline systems using stochastic optimization and GIS Ttools. *Chem. Eng. Res. Des.* **2012**, *90*, 2209.
- (30) Kirkpatrick, S. Optimization by simulated annealing: Quantitative studies. *J. Stat. Phys.* **1984**, *34*, 975.
- (31) Brooke, A.; Kendrick, D.; Meeraus, A.; Raman, R. *GAMS: A Users Guide*; GAMS Development Corp.: Washington, DC, 2004.
- (32) Ulrich, G. D.; Vasudevan, P. T. *Chemical Engineering Process Design and Economics*, 2nd ed.; Process Publishing, 2004; Ch. 5.
- (33) Winston, W. L. *Opera Research: Applications and Algorithms*, 4th ed.; Brooks/Cole Duxbury Press, 2004; Ch. 7.

- dx.doi.org/10.1021/ie400022g | *Ind. Eng. Chem. Res.* 2013, 52, 8579–8588